



(11)

EP 1 231 276 A1

(12)

EUROPEAN PATENT APPLICATION

(43) Date of publication
14.08.2002 Bulletin 2002/33

(51) Int. Cl. **C12N 15/74, C12N 15/11,
C12N 15/62, C12P 21/02**

(21) Application number: **02004765.0**

(22) Date of filing: **17.08.1999**

(84) Designated Contracting States
**AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU
MC NL PT SE**

• **Rigaut, Guillaume**
69007 Lyon (FR)

(30) Priority: **17.08.1998 EP 98115448**

(74) Representative: **Dey, Michael, Dr. et al**
Weickmann & Weickmann,
Patentanwälte,
Postfach 86 08 20
81635 München (DE)

(62) Document number(s) of the earlier application(s) in
accordance with Art. 176 EPC
99944420.1 / 1 105 508

(71) Applicant: **EUROPÄISCHES LABORATORIUM
FÜR MOLEKULARBIOLOGIE (EMBL)**
D-69117 Heidelberg (DE)

Remarks

This application was filed on 01.10.2002 as a
divisional application to the application mentioned
under IN ID code 62.

(72) Inventors

• **Seraphin, Bertrand**
78114 Magny les Hameaux (FR)

(54) **Method for purifying proteins and/or biomolecule of protein complexes**

(57) The present invention relates to a method for detecting and/or purifying proteins and/or biomolecule or protein complexes as well as fusion proteins, nucleic acids, vectors and cells suitable for this method.

EP 1 231 276 A1

Description

[0001] The present invention relates to a method for purifying substances, such as biomolecules, proteins, protein and/or biomolecule complexes, subunits of biomolecule complexes, or components, or components or even whole cells, of a selected fusion proteins for use in this method and other related subjects.

[0002] Protein expression and purification methods are essential for studying the structure, activities, interactions with other proteins, induced effects of proteins of interest. Methods that are currently available use systems such as bacteria or cells transfected with expression vectors or infected with baculovirus.

[0003] In order to study individual proteins a first requirement is to obtain sufficient amounts of that particular protein to be able to carry out biological and biochemical analyses such as activity tests, interaction assays, structure determination and the like. For this purpose the genes coding for the proteins of interest are cloned into vectors that allow the expression of those proteins in suitable host cells, usually the proteins are expressed at high levels. This overexpression leads to the generation of large amounts of protein but often has the disadvantage of yielding insoluble protein which is present in so called inclusion bodies in the cells. The overexpressed protein then has to be resolubilized before analysis. Although such methods work well for conventional protein detection methods based on weight analysis (polyacrylamide gels, Western blots, etc.) of the expressed protein, they are not suitable for other studies and for assays on protein complexes.

[0004] Currently used protocols for over-expressing proteins are normally carried out in host cells in which the protein of interest is not normally expressed (for example when eukaryotic proteins are expressed in bacteria). Apart from the insolubility problems expressed in this manner often show a lack of proper post-translational or post-translational modification such as correct processing or glycosylation.

[0005] Another disadvantage is that the expression of one subunit of a complex, even in a homogeneous system, might not lead to increased production of all the complex components and in some extreme cases can even result in misfolding of that protein of interest to an aberrant complex (e.g. the proteasome, Swafford et al. 1996, Nature 379, 658).

[0006] Purification of proteins expressed at the high level is therefore indispensable in many cases but appropriate purification protocols are not easily available because of the huge biochemical knowledge required to establish a suitable protocol. Developing a purification scheme requires assessing the chemical and physical properties of the target protein by a trial and error process, making it tedious and time consuming because such analyses have to be repeated for each protein.

[0007] Affinity purification methods for the purification

of proteins are known. Kits and apparatuses for affinity chromatography are commercially available. Usually a fusion protein of a polypeptide of interest plus an affinity tag is expressed in a system such as one of the above mentioned expression systems, the fusion protein is extracted and added to support material, such as a resin or matrix packed in an affinity column which is loaded with a material that specifically binds the affinity tag. After binding of the target protein to the matrix via the affinity tag, unbound substances can be removed simply by washing the matrix. The fusion protein can then be eluted off the matrix using chemical agents or specific temperature or pH conditions. Since the affinity tags usually bind with high affinity stringent conditions are needed for the subsequent elution which can often destroy/damage or denature the protein of interest.

[0008] Affinity tags possess groups or moieties which are capable of binding to a specific binding partner with high affinity. Various affinity tags are known in the art and have been widely used for the purification of proteins. Examples are the ligand binding domains of protein A of *Staphylococcus aureus*, glutathione S-transferase (GST), maltose binding protein, cellulose binding domain, polyarginine, polylysine, polyhistidine, polyphenylalanine, leucine zipper, leucine binding domains. These bind with high affinity to an appropriate matrix which is covered with the specific binding partner. In the case of protein A, IgG coated Sepharose has been used for affinity chromatography of fusion proteins possessing a protein A domain (Senger et al. 1996, EMBO J. Vol. 17, 2196-2207). Other examples are discussed in Sassenfeld, TIBTECH, 1490, p. 88. A plasmid vector containing a cassette encoding a camouflin binding peptide is available from Stratagene.

Normally fusion proteins are tagged with only one affinity tag and are purified in a single purification step. This often leads to problems due to remaining contaminants. Another limitation of most of the conventional methods is that they are adapted for expression of the proteins in bacteria only. WO96/40443 discloses a method of expressing fusion proteins in gram positive bacteria either anchored to the membrane or in secreted form. The anchored proteins are cleaved off using TEV protease and subsequently affinity purified via an affinity tag.

[0009] Often the affinity tag is removed from the fusion protein after the affinity purification step by the action of a specific protease, such as the TEV protease. This means however that the purified fraction contains substantial amounts of this protease (Senger et al. 1996) which severely limits the applications of such protein preparations.

[0010] It is therefore an object of the present invention to provide a purification/detection method for proteins and/or biomolecule or protein complexes and/or components or subunits thereof which eliminates the disadvantages of the currently known methods and allows efficient purification not only of affinity tagged target proteins as fusion proteins but also of other substances that

are capable of binding with these proteins.

One method according to the invention for purifying substances selected from proteins, biomolecules, complexes of proteins or biomolecules, subunits thereof or their components, oligonucleotides, and whole cells, such as those in the following steps:

(a) providing an expression environment containing one or more heterologous nucleic acids encoding one or more polypeptides and/or one or more subunits of a biomolecule complex, the polypeptides or subunits being fused to at least two different affinity tags, one of which consists of one or more IgG binding domains of Staphylococcus protein A;

(b) maintaining the expression environment under conditions that facilitate expression of the one or more polypeptides or subunits in a native form as fusion proteins with the affinity tags;

(c) detecting and/or purifying the one or more polypeptides or subunits by a combination of at least two different affinity purification steps, each comprising binding the one or more polypeptides or subunits via one affinity tag to a support material capable of selectively binding one of the affinity tags and separating the one or more polypeptides or subunits from the support material after substances not bound to the support material have been removed.

[0011] An alternative method of the invention which is particularly suitable for detecting and/or purifying protein or biomolecule complexes is a method comprising the steps:

(a) providing an expression environment containing one or more heterologous nucleic acids encoding at least two subunits of a biomolecule complex, each being fused to at least one of different affinity tags, one of which consists of one or more IgG binding domains of Staphylococcus protein A;

(b) maintaining the expression environment under conditions that facilitate expression of the one or more subunits in a native form as fusion proteins with the affinity tags and under conditions that allow the formation of a complex between the one or more subunits and possibly other components capable of complexing with the one or more subunits;

(c) detecting and/or purifying the complex by a combination of at least two different affinity purification steps, each comprising binding the one or more subunits via one affinity tag to a support material capable of selectively binding one of the affinity tags and separating the complex from the support material after substances not bound to the support material have been removed.

[0012] For the purpose of the invention, a biomolecule can be a protein, polypeptide or a nucleic acid or other

biomolecule. A biomolecule complex denotes a complex of at least two biomolecules, preferably at least one protein associated either with other proteins which are then called subunits or with other substances which can for example be nucleic acids. The biomolecule complexes can be natural ones, such as nuclear ribonucleic acid (rRNA) and ribulose biphosphate complex, or they can be artificial ones, such as mutant DNA binding proteins associated with mutant target DNAs. Any complex molecule comprising one or more subunits is a polypeptide or a subunit expressed according to the invention and/or further comprising other components which associate in a manner stable enough not to be dissociated by the affinity steps. A biomolecule complex that can be detected and/or purified by the method of the invention. A protein complex generally denotes a complex between protein subunits.

[0013] The nucleic acid sequence of the protein to be purified must be known or at least available so that it can be cloned into a nucleic acid which is suitable to drive expression in the appropriate host cells or cell-free expression systems. If a protein complex is to be purified, the nucleic acid sequence of at least one of its subunits has to be known or available.

[0014] The heterologous nucleic acid driving the expression of the protein to be purified according to the invention thus contains appropriate sequences that allow it to be maintained in the chosen host cell or cell-free system, such as a promoter and, if necessary, other control sequences such as enhancers and poly A's tail.

[0015] In principle any host cell that is competent to take the heterologous nucleic acid from which the polypeptides or subunits are to be expressed is suitable as an expression environment. These cells can be prokaryotic cells such as bacteria etc. or eukaryotic cells such as yeast, fungi or mammalian cells. Preferably the protein or subunit or protein complex to be purified is expressed in its natural host. Since this is, where is very efficient, the proteins are preferably expressed at their natural levels. This has the advantage of avoiding the formation of inclusion bodies and also reduces the risk of toxic effects on the cell that large amounts of certain proteins may have. Furthermore, this avoids purifying excess protein subunits that are not assembled into a complex or that are assembled into aberrant complexes (see above).

[0016] After the heterologous nucleic acid encoding the fusion protein has been introduced into a chosen host cell, the cells are cultured under conditions which allow the expression of the fusion protein(s).

[0017] As a ready method for controlling transcriptional control sequences are preferably selected so that the fusion protein is not over-expressed but is expressed at basal levels in the cell. This serves to ensure that the protein is expressed in a native form. Native form means, in this context, that a conformationally close to natural three-dimensional structure of the protein is achieved, i.e. the protein is folded correctly. More preferably, the protein will also be processed correctly and show normal post-

transcriptional and post-translational modification. The correct folding is of great importance, especially when the expressed polypeptide is a subunit of a protein complex because it will bind to the other subunits of the complex only when it is present in its native form. However, this also poses problems for expressing mutant proteins. These can also have a native conformation. Such mutant subunits can, for example, be used to purify mutant complexes or complexes that contain one or other mutated subunits.

[0018] Depending on the protein or subunit to be purified, the fusion protein is expressed intracellularly or secreted into the culture medium. Alternatively, it might be targeted to other cell compartments such as the membrane. Depending on the protein an appropriate method is used to extract the fusion protein from the cells and/or medium. When a fusion protein is expressed which is targeted to a certain subcellular location, e.g. the membrane of cell organelles or the cell membrane, these organelles or the cells themselves can be purified via the binding of these membrane proteins. It is also possible to purify cell organelles via proteins naturally expressed on their surface which bind to the fusion protein of the invention.

[0019] Further, it is possible to purify biomolecule or protein complexes, subunits or other substances that are capable of binding to or complexing with the fusion protein generated according to the invention. These substances can bind to fusion protein either directly or via linker/modulators. Linker/modulators in this context may be anything which is capable of binding two or more biomolecules so that these biomolecules are then part of a complex although they may not be directly associated with each other.

[0020] According to the invention, it is also possible to use cell-free systems for the expression of the polypeptides or subunits. These must provide all the components necessary to effect expression of proteins from the nucleic acid such as transcription factors, enzymes, ribosomes etc. in vitro transcription and translation systems are commercially available as kits so that it is not necessary to describe these systems in detail (e.g. rabbit reticulocyte lysate systems for translation). A cell-free or in vitro system should also allow the formation of complexes.

[0021] For the purification according to the invention, it is preferable to employ affinity chromatography or affinity columns when carrying out a matrix loaded with the appropriate binding partner for the affinity tag used in that particular purification step.

[0022] In accordance with the method of this invention, two affinity steps are carried out. Basically each affinity step consists of a binding step in which the previously extracted protein is bound via one of its affinity tags to a support material which is covered with the appropriate binding partner for that affinity tag. Then unbound substances are removed and finally the protein to be purified is recovered from the support material. This can be

done in two ways. The first possibility is to simply use conventional elution techniques such as varying the pH or the salt or buffer concentrations and the like depending on the tag used. The second possibility is to release the protein to be purified from the support material by proteolytic cleavage of the affinity tag bound to the support. In this way, the protein can be recovered in the form of a truncated fusion protein or "free" affinity tags have been cleaved off as the target polypeptide or subunit itself.

[0023] According to one embodiment of the present invention a fusion protein of a single polypeptide plus two different affinity tags is expressed, wherein one of the tags comprises one or more IgG binding domains of protein A of *Staphylococcus aureus*.

[0024] More preferably, a specific proteolytic cleavage site is present in the fusion protein between the one or more polypeptides or subunits and the one or more affinity tags so that proteolytic cleavage allows the removal of at least one of the affinity tags, especially the IgG binding domains of protein A.

[0025] Proteolytic cleavage can be carried out by chemical means or enzymatically.

The proteolytic cleavage site that is used to cleave off one of the affinity tags is preferably an enzymatic cleavage site. There are several proteases which are highly specific for short amino acid sequences which they will cleave. One of these is a specific cleavage site of Tobacco Etch Virus (TEV) which is cleaved by the TEV protease N.A. Recombinant TEV protease is available from Gibco BRL. The TEV cleavage site is preferably used as the cleavage site to remove the protein A domains from the fusion protein.

[0026] Even more preferably, the affinity step using protein A binding to IgG is carried out first by binding the one or more polypeptides or subunits via the one or more IgG binding domains of *Staphylococcus* to a support material capable of specifically binding the latter, removing substances not bound to the support material and separating the one or more polypeptides or subunits from the support material by cleaving off the IgG binding domains via the specific proteolytic cleavage site, and then another affinity tag is used to purify the protein further via a conventional elution step comprising binding the polypeptide or subunit via another affinity tag to a second support material capable of specifically binding the latter, removing substances not bound to the support material and separating the polypeptide or subunit from the support material.

[0027] When the proteins are present at low concentrations in the expression environment and/or the support material, a large amount of protease is required to release the bound material from the support material, in other words, when the substrate concentration is low a high level of enzyme is required to drive an efficient proteolytic reaction. The second purification step is then important to remove remaining contaminants and the protease. Removal of the protease is preferable in order to

ent provide any negative influence of the proteolytic activity on the purification.

[0028] However, in some cases it may be desirable to remove all the affinity tags, in which case it is also possible to use two or more different proteolytic cleavage sites for the separation of the polypeptides/subunits from the support material.

[0029] The method according to the invention naturally facilitates efficient purification of proteins of interest but also allows "fishing" for and detecting components present in complexes with which the polypeptides or subunits are associated or complexed either directly or indirectly, e.g. molecules such as linker molecules. This would allow selective fishing for certain substances which may be potential drugs even from complex mixtures.

[0030] According to a second embodiment of the invention it is possible not only to detect or purify the subunit containing fusion proteins expressed but also other subunits that are capable of associating with the proteins expressed in a direct way, i.e. by directly binding to the fusion protein, or indirectly via other molecules to form complexed complexes. If a fusion protein of a subunit of a biomolecule or protein complex is purified according to the invention the affinity steps are chosen so that other complex components which have bound to the fusion protein are still associated with the subunit after the purification steps so that they can be detected, purified as well.

[0031] The biomolecule complexes can be formed in the expression environment such as cellular complexes. Alternatively, other complex components may be added to the subunits already expressed to form complexes in vitro or may even be added when the subunit containing fusion protein is already bound to a support material in an affinity step.

[0032] It is also possible to express two or more subunits of the same complex each as a fusion protein with a different affinity tag. When the subunits associated they can be detected/purified possibly together with other complex components by a series of affinity steps in which each time the complex is bound via a differently tagged subunit. The two or more affinity tags can be fused with a single subunit of a complex or with two or more subunits which bind to each other or are simply present in the same complex.

[0033] The purification steps can be carried out as described above.

[0034] Some polypeptides are present in more than one complex so that the components of all complexes can be purified.

[0035] If one is interested in a single complex A one can also subtract other complexes B that also contain one of the subunits of A by fusing that subunit of A to one tag and fusing a subunit unique to B with a different tag. The tagged subunit of B will bind to a specific support material. If the fusion first bound to that support material is used in the second affinity purification step

complex B will no longer be present because it was removed, subtracted, in the first affinity step. Many similar scenarios can be envisaged and designed by a person skilled in the art.

[0036] Further affinity tags in addition to the IgG binding domains that can be used in accordance with the present invention can be any conventional affinity tag. Preferably the second affinity tag consists of at least one calcium binding peptide (CBP). Calcium binding peptide as an affinity tag has been described and is commercially available (Stratagene). When a calcium binding peptide is used the corresponding purification steps are carried out using a support material that is coated with calcium. The calcium binding peptide tag binds to calcium in the presence of low concentrations of calcium. It can subsequently be eluted using a chemical agent such as a chelating agent like EDTA. Preferably around 2 mM EDTA is added for the elution step.

[0037] Another aspect of the present invention is a fusion protein consisting of one or more polypeptides or subunits fused to at least two affinity tags, wherein one of the affinity tags consists of at least one IgG binding domain of Staphylococcus protein A.

[0038] Other fusion proteins according to the invention are those additionally including a proteolytic cleavage site, preferably to cleave off the IgG binding domains, or those in which the second tag represents one or more CBPs. Again, the skilled person will be able to select and construct the most suitable combinations of tags and cleavage sites and polypeptides and/or subunits in fusion proteins depending on the affinity strategy used. The fusion protein can be constructed so that the above-mentioned purification, detection or fishing procedures can be carried out.

[0039] There are several possibilities for constructing the fusion protein. In principle, the affinity tags may be fused close to either of the N- or C-terminal ends of the polypeptides or subunits to be expressed. The order in which the tags are fused with the polypeptides or subunits is, of course, not set out but can be chosen according to the affinity protocol to be used. Small peptides such as the CBP can even be fused to the polypeptides of interest internally, as long as the reading frame on the nucleic acid is not disrupted. Preferably the tags are located near to the same end of the polypeptides or subunits, wherein it is especially preferred that the IgG binding domains are placed at the N- or C-terminus of the complete fusion protein, followed by a proteolytic cleavage site, the other tag(s) and the polypeptides or subunits.

[0040] The fusion protein can also contain a second proteolytic cleavage site for the removal of the second affinity tag. The most preferable combination of affinity tags and cleavage sites is the one with protein A domains of Staphylococcus as the first affinity tag which can be cleaved off via the TEV protease and using at least one calcium binding peptide as the second af-

affinity tag which allows the isolation of the purified fusion protein using a chelating agent such as 2 mM EDTA.

[0041] Another aspect of the present invention is a heterologous nucleic acid coding for a fusion protein as the one described above.

[0042] A further aspect of the invention is a vector comprising at least one heterologous nucleic acid coding for a fusion protein of the invention. This vector contains the nucleic acid under the control of sequences which facilitate the expression of the fusion protein in a particular host cell or cell-free system. The control sequences comprise sequences such as promoters and, if necessary, enhancers, poly A sites etc. The promoter and other control sequences are selected so that the fusion protein is preferably expressed at a basic level so that it is produced in a soluble form and not as an insoluble material. Preferably the fusion protein is also expressed in such a way as to allow correct folding for the protein to obtain a native conformation. Preferably, one or more selectable markers are also present on the vector for the maintenance in prokaryotic or eukaryotic cells. Basic cloning vectors are described in Sambrook et al., Molecular Cloning, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, (1989). Examples of vectors are plasmids, bacteriophages, other viral vectors and the like.

[0043] In a preferred embodiment vectors are constructed containing pre-made cassettes of affinity tag combinations into which the nucleic acid coding for the polypeptide or subunit of interest can be inserted by means of a multiple cloning site such a polynucleotide linker. Thus, a vector according to the invention is also one which does not contain the coding sequences for the polypeptide(s) or subunit(s) of interest but contains the above-mentioned components plus one or more polynucleotide linkers with preferably unique restriction sites in such a way that the insertion of nucleic acid sequences according to conventional cloning methods into one of the sites in the polynucleotide linker leads to a vector encoding a fusion protein of the invention.

[0044] In a further preferred embodiment the vector comprises heterologous nucleic acid sequences in form of two or more cassettes each comprising at least one of different affinity tags, one consisting of one or more IgG binding domains of Staphylococcus aureus protein A and at least one polynucleotide linker for the insertion of further nucleic acids. Such a vector can be used to express two subunits of a protein complex, each tagged with a different tag.

[0045] Vectors according to the invention can be introduced into host cells stably or transiently, they can be present extrachromosomally or integrated into the host genome, and they can be used to produce recombinant cells or organisms such as transgenic animals.

[0046] Another object of the invention is a cell containing a heterologous nucleic acid or a vector of the invention. These cells can be prokaryotic or eukaryotic cells, e.g. bacteria, cells, yeast cells, fungi, or mammalian

cells, and the vector or nucleic acid can be introduced (transformed) into these cells stably or transiently by conventional methods, protocols for which can be found in Sambrook et al. *supra*.

[0047] Yet a further aspect of the invention is a reagent kit, preferably comprising vectors as described above together with support materials for carrying out the affinity steps. The support materials carry moieties which are capable of specifically binding the affinity tags, for example, calmodulin coated resin in the case of calmodulin binding peptide as the affinity tag or IgG-coated resins for affinity tags consisting of protein A domains. Additionally such a kit may comprise buffers and other conventional materials for protein purification, especially for affinity chromatography. Further, the kit preferably provides at least one proteolytic agent such as a chemical agent capable of performing proteolysis or a protease, and/or chemical agents such as oxidizing agents where the protease is capable of proteolytically cleaving the fusion protein. When two proteolytic cleavage sites are used the kit will preferably contain two proteases.

[0048] The following examples and figures serve to illustrate the invention and its practical application; they are, however, not intended to limit the scope of the invention.

Fig. 1 shows a Coomassie stained gel depicting the fractionated proteins of a yeast RNA-protein complex. Proteins identified by mass spectrometry are labeled 1-24. Bands 1, 3, 8-11, 16-24 were expected in this complex. Bands 2, 4-7 represent proteins that are likely bandmates for true complex component given their sequence. Bands 12-14 represent potential contaminants (ribosomal proteins). Band 15 is a trace amount of T6V protease that is introduced in this particular preparation. This is not generally the case, see Fig. 2 for example. Bands 4, 6 originate from the same gene and might represent alternative translation products or degradation products. Band 16 is a mixture and contains in addition to a bona fide complex protein a contaminating ribosomal protein.

Fig. 2 shows a Coomassie stained gel depicting the fractionated proteins moiety of the yeast U1 snRNP. The 10 specific proteins were identified by mass spectrometry. Compared with the silver stained gel data had following a purification using a cap affinity step and a NEA and unit reported by Neubauer et al., Proc Natl. Acad Sci USA 1997, 94, 355-359. Note in particular the low level of contaminants in this purification.

Fig. 3 shows a Coomassie stained gel in which purified U1 snRNP has been analysed using the 3B4 binding E6TA column and a band 1

the Protein A binding TEV cleavage and 3) or both steps (and 7). Arrows on the right point to the U1 snRNP specific proteins. Lanes 2, 4 and 8 show the background signals obtained using each of the single steps or the two step procedure from an extract without tagged protein, demonstrating the need for two steps to get a pure material. Lane 6 is a molecular weight marker. Lane 7 is the TEV protease that can also be seen as an abundant contaminant (even though only the required amount of protease was used in lanes 3 and 4). This demonstrates again the need for a second purification step.

Examples

Example 1

Purification of protein complexes from yeast.

[0049] A vector encoding a fusion of a yeast protein to the CBP-TEV-Protein A double tag was constructed using standard methods. The fusion protein is one subunit of a protein complex of yeast containing 24 subunits in total. The plasmid was transformed into yeast cells and a 2 l culture of cells expressing the protein was prepared. Proteins were extracted from the cultured cells using a French press. The complex was purified by binding to IgG-linked beads, eluting by TEV protease cleavage, binding of the eluted material on carmodulin containing beads followed by elution with EGTA. All steps were carried out at 0-4 °C, excepted for TEV cleavage.

[0050] The first affinity step (IgG step) was performed as follows:

200 μ l of IgG-Sepharose bead suspension (Pharmacia L709490) were washed in an Econocolumn with 5 ml of RPR-150 IgG buffer (10 mM TrisCl, pH 8.0, 150 mM NaCl, 0.1 mM NP40). 10 ml extract (corresponding approximately to 2 l of yeast culture) were adjusted to RPR-150 IgG buffer concentrations (10 mM TrisCl, pH 8.0, NaCl and NP40). This extract solution was mixed with the 200 μ l of IgG-Sepharose beads and rotated in the Econocolumn for 2 hours. The unbound fraction was discarded and beads with bound material were washed first with 30 ml RPR-150 IgG buffer followed by 10 ml TEV cleavage buffer (10 mM TrisCl, pH 8.0, 150 mM NaCl, 0.1 mM NP40, 0.5 mM EDTA, 1 mM DTT).

[0051] The target protein was cleaved and released from the beads as follows. The washed Econocolumn was filled with 1 ml TEV cleavage buffer and 30 μ l TEV protease and rotated in a 37 °C incubator for 2 hours. The eluate was recovered by gravity flow.

[0052] The second affinity step (Carmodulin affinity step) was performed as follows:

The previous eluate was mixed with 3 ml of RPR-150 Carmodulin binding buffer (10 mM TrisCl, pH 8.0, 150 mM NaCl, 1 mM Mg acetate, 1

mM imidazole, 2 mM CaCl₂, 0.1 mM NP40). The appropriate amount of CaCl₂ was further added to block the EDTA coming from the TEV cleavage buffer. This mix was rotated for 1 hour in an Econocolumn containing 200 μ l of Carmodulin beads 8 μ ny (Stratagene 21430) previously washed with 5 ml RPR-150 Carmodulin binding buffer.

[0053] After washing with 30 ml of RPR-150 Carmodulin binding buffer, protein complexes were eluted with successive additions of 200 μ l of RPR-150 Carmodulin elution buffer (10 mM β -mercaptoethanol, 10 mM TrisCl, pH 8.0, 150 mM NaCl, 1 mM Mg acetate, 1 mM imidazole, 2 mM EGTA, 0.1 mM NP40).

[0054] Samples were frozen in dry ice and stored at -80 °C. Proteins were concentrated by TCA precipitation (A. Bensadoun and D. Weinstein, 1976; Anal. Biochem. 70: 241). The proteins were detected by polyacrylamide gel electrophoresis with subsequent staining of the gel with Coomassie blue. The result of the protein purification (a gel of which is depicted in Fig. 1) demonstrates that the strategy employed is highly efficient. All the expected protein subunits which number 24 in this case can be detected.

Example 2

[0055] The same procedure was used for two other protein or protein RPA complexes from yeast where all expected protein subunits were detected using the method of the invention. These are the CBC (Cap Binding Complex) and the U1 snRNP. The purified U1 snRNP is depicted in Figures 2 and 3. The CBC complex has been shown to be still active and the purity was good in all cases. This method is relatively cheap and not very time-consuming, since it can be done in one day. Concerning the U1 snRNP, it is noteworthy that Neuhauer et al. (Proc. Natl. Acad. Sci. USA 1997, 94: 385-390) carried out a purification of the same complex (U1 snRNP) extracted from 18 l of culture. The proteins of the complex of interest were then only visible by silver staining and several contaminants were still observed.

Claims

1. Method for detecting and/or purifying substances selected from proteins, nucleic acids, cell components, cell organelles and cells comprising the steps:

a) providing an expression environment containing one or more heterologous nucleic acids encoding one or more polypeptides and/or one or more subunits of a biomolecular complex, the polypeptides or subunits being fused to at least two different affinity tags, one of which consists of one or more IgG binding domains of Staphylococcus protein A.

- for maintaining the expression or refinement of our conditions that facilitate expression of the one or more polypeptides or subunits in a native form as fusion proteins with the affinity tags.
- (c) detecting and/or purifying the one or more polypeptides or subunits by a combination of at least two different affinity purification steps, each comprising binding the one or more polypeptides or subunits via one affinity tag to a support material capable of selectively binding one of the affinity tags and separating the one or more polypeptides or subunits from the support material after substances not bound to the support material have been removed.
2. Method according to claim 1, wherein between the one or more polypeptides or subunits and one or more of the affinity tags a specific proteolytic cleavage site is present in the fusion protein which facilitates the removal of one or more of the affinity tags.
 3. Method according to claim 2, wherein the specific proteolytic cleavage site is an enzymatic cleavage site.
 4. Method according to claim 3, wherein the specific proteolytic cleavage site is the cleavage site for TEV protease N/A.
 5. Method according to claim 2, 3 or 4, wherein the proteolytic cleavage site is used to cleave the polypeptide or subunit in step (c) from the IgG binding domain of Staphylococcus protein A bound to the support material.
 6. Method according to claim 5, wherein the affinity purification of step (c) comprises:
 - (i) binding the one or more polypeptides or subunits via the one or more IgG binding domains of Staphylococcus to a support material capable of specifically binding the latter, removing substances not bound to the support material and separating the one or more polypeptides or subunits from the support material by cleaving off the IgG binding domains via the specific proteolytic cleavage site; and
 - (ii) rebinding the polypeptide or subunit via another affinity tag to a second support material capable of specifically binding the latter, removing substances not bound to the support material and separating the polypeptide or subunit from the support material.
 7. Method according to claim 6, wherein step (i) is carried out before step (ii).
 8. Method according to one of the previous claims, wherein the fusion protein contains a second specific proteolytic cleavage site for the removal of one or more of the other affinity tags.
 9. Method according to one of the previous claims, wherein one of the affinity tags consists of at least one IgG binding domain.
 10. Method according to claim 9, wherein a chemical agent is used to separate the one or more polypeptides or subunits from the support material.
 11. Fusion protein comprising at least one polypeptide or subunit of a protein complex fused to at least two different affinity tags, wherein one of the affinity tags consists of at least one IgG binding domain of Staphylococcus protein A.
 12. Fusion protein according to claim 11, wherein the protein additionally contains a specific proteolytic cleavage site.
 13. Nucleic acid coding for a fusion protein according to claim 11 or 12.
 14. Vector comprising a nucleic acid according to claim 13 under the control of sequences facilitating the expression of a fusion protein according to claim 12 or 13.
 15. Vector comprising heterologous nucleic acid sequences in form of one or more cassettes, each comprising at least two different affinity tags, one consisting of one or more IgG binding domains of Staphylococcus aureus protein A, and at least one polynucleotide linker for the insertion of further nucleic acids.
 16. Vector comprising heterologous nucleic acid sequences in form of two or more cassettes, each comprising at least one of the different affinity tags, one consisting of one or more IgG binding domains of Staphylococcus aureus protein A, and at least one polynucleotide linker for the insertion of further nucleic acids.
 17. Cell containing a nucleic acid according to claim 13 or a vector according to claim 14.
 18. Reagent kit comprising a nucleic acid according to claim 13 or a vector according to claim 14, 15 or 16 for the expression of a fusion protein according to claim 11 or 12 and support material(s) each capable of specifically binding one of the affinity tags.
 19. Reagent kit according to claim 18 additionally comprising at least one chemical agent for separating one of the affinity tags from its support material and/or

of a specific chemical probe, lyso agent and/or sonication protease capable of cleaving the fusion protein.

20. Use of the method according to one of claims 1 to 19 for the detection and/or purification of substances capable of interacting with the fusion protein.
21. Use of the method according to one of claims 1 to 19 for the detection and/or purification of cells and/or cell organelles expressing the fusion protein on their surface.

15

20

25

30

35

40

45

50

55



European Patent
Office

EUROPEAN SEARCH REPORT

Application Number
EP 02 00 4761

DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int. Cl. 7)
X	WO 96 40943 A (UNIV. ROCKEFELLER; DARZINS ALDIS (US); WHITEHEAD STEPHEN (US); HRUB) 19 December 1996 (1996-12-19) * abstract; claims 1-21; figure 6; examples 1-3 * * page 9 - line 10 - line 19 *	1-4, 1-15, 17-19, 5-7,9, 10,20	C12N15/74 C12N15/11 C12N15/62 C12P21/02
Y, D	SENGER B ET AL: "Mtr10p functions as a nuclear import receptor for the mRNA binding protein Npl3p" EMBO JOURNAL, vol. 17, no. 8, 1 August 1998 (1998-08-01), pages 2196-2207, XP002090021 * the whole document *	5-7,20	
Y	ZHENG C ET AL: "A new expression vector for high level protein production, one step purification and direct isotopic labelling of calmodulin-binding peptide fusion proteins" GENE, vol. 186, no. 1, 20 February 1997 (1997-02-20), page 55-60, XP004054879 * the whole document *	9,10	
			TECHNICAL FIELDS SEARCHED (Int. Cl. 7)
			C12N C12P
A	PANAGIOTIDIS C A ET AL: "pALEX, a dual-tag prokaryotic expression vector for the purification of full-length proteins" GENE, vol. 164, no. 1, 1995, page 45-47, XP004041915 * the whole document *		

The present search report has been drawn up for all claims

THE HAGUE

19 June 2002

Oderwald, H

**ANNEX TO THE EUROPEAN SEARCH REPORT
ON EUROPEAN PATENT APPLICATION NO.**

EP 02 00 4765

The following table lists the patent family members searched in the context of the present European patent application.
The table is based on data as contained in the European Patent Office (EPO) database.
The European Patent Office (EPO) was not able to trace prior art which is directly given in the present information.

19-06-2002

Patent document used in search report	Publication date	Patent family member	Publication date
WO 9540943 A	19-12-1995	US 5821088 A	13-10-1998
		AJ 711985 B2	28-10-1999
		AJ 6331096 A	30-12-1996
		CA 2224172 A1	19-12-1996
		CN 1192245 A	02-09-1998
		EP 0832252 A1	01-04-1998
		FI 974432 A	02-02-1998
		HU 9902004 A2	28-10-1999
		JP 2000512121 T	19-09-2000
		NZ 311902 A	23-12-1998
		WO 9540943 A1	19-12-1995

For more details about this annex, see Official Journal of the European Patent Office (OJEP).